Computers & Graphics 40 (2014) 10–21

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Technical Section

Immersive full-surround multi-user system design ☆

JoAnn Kuchera-Morin [a,*], Matthew Wright [a], Graham Wakefield [a], Charles Roberts [a], Dennis Adderton [a], Behzad Sajadi [b], Tobias Höllerer [a], Aditi Majumder [b]

[a] AlloSphere Research Group and Media Arts Technology Program, University of California Santa Barbara, United States
[b] Department of Computer Science, University of California Irvine, United States

A R T I C L E I N F O

Article history:
Received 3 April 2013
Received in revised form
6 December 2013
Accepted 20 December 2013
Available online 18 January 2014

Keywords:
VR systems
Display technology
Multi-user
Multimodal interaction
Immersion

A B S T R A C T

This paper describes our research in full-surround, multimodal, multi-user, immersive instrument design in a large VR instrument. The three-story instrument, designed for large-scale, multimodal representation of complex and potentially high-dimensional information, specifically focuses on multi-user participation by facilitating interdisciplinary teams of co-located researchers in exploring complex information through interactive visual and aural displays in a full-surround, immersive environment. We recently achieved several milestones in the instrument's design that improves multi-user participation when exploring complex data representations and scientific simulations. These milestones include affordances for "ensemble-style" interaction allowing groups of participants to see, hear, and explore data as a team using our multi-user tracking and interaction systems; separate visual display modes for rectangular legacy content and for seamless surround-view stereoscopic projection using 4 high-resolution, high-lumen projectors with hardware warping and blending integrated with 22 small-footprint projectors placed above and below the instrument's walkway; and a 3D spatial audio system enabling a variety of sound spatialization techniques. These facilities can be accessed and controlled by a multimodal framework for authoring applications integrating visual, audio, and interactive elements. We report on the achieved instrument design.

1. Introduction

This paper presents design decisions and results from 5 years of ongoing research involving the AlloSphere [1,2], a three-story, immersive instrument designed to support collaborative scientific/artistic data exploration and empower human perception and action. To support group experiences of research, working, and learning, we believe that computer systems need to accommodate physically co-located users in immersive multimodal[1] environments. We focus on research driving the full-surround, immersive, and multimodal aspects of the facility, allowing content to drive its technological development. Research in the facility is thus two-fold: (1) multimedia systems design to develop a large, interactive, multimodal instrument, and (2) data generation, representation, and transformation – using a diverse set of applications to drive the development of the instrument's capabilities for real-time interactive exploration. Our research maxim is that content drives technology, with no feature being added to our production system without first being explored in a prototype application. Our facility is designed to operate in two modes: *desktop* mode provides the opportunity to bring legacy content quickly into the system for rapid turnaround, while *surround* mode facilitates full-surround immersion (as shown in Fig. 1).

We believe that interdisciplinary teams encompassing the physical sciences, life sciences, social sciences as well as the arts will produce audiovisual data representations that will lead to increased understanding of large and complex biological systems, social networks, and other heterogeneous, high-dimensional information. The design process for our instrument and its computational infrastructure has thus been driven by the goal of providing multi-user capabilities supporting interdisciplinary research teams.

We designed, built, and equipped our facility using in-house planning and expertise, rather than relying on a commercial or integrator-driven solution. The physical infrastructure includes a large perforated-aluminum capsule-shaped screen (two 16-foot-radius tilt-dome hemispheres connected by a 7-foot wide cylindrical section) in a three story near-to-anechoic room. A 7-foot-wide bridge through the center of the facility provides space for up to 30 users simultaneously. The hemispheres' locations on the sides instead of overhead and underneath support the concept of looking to the horizon at the equator of the instrument's

☆This article was recommended for publication by Dirk Reiners.

* Corresponding author. Tel.: +1 805 893 3010, Mobile: +1 805 729 2272, fax: +1 805 893 2930.
E-mail address: jkm@create.ucsb.edu (J. Kuchera-Morin).

[1] Here by "multimodal" we are specifically referring to vision, hearing, and physical interaction.

0097-8493/$ - see front matter Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.cag.2013.12.004

Fig. 1. Fisheye photographs of multiple users interacting with full-surround audiovisual content in real time. AlloBrain (left), ray-traced cuboids (center), and world map (right).

infrastructure, while the joining cylindrical section avoids the in-phase acoustic echoes that would be present inside a perfectly spherical structure. The perforated screen allows for the 3D spatial audio system as well as the multi-user tracking system to be placed outside the sphere.

Over the past few years, we have focused on true multimodality, attempting an equal balance among visual, audio and interactive representation, transformation and generation across a diverse set of content areas. We have also concentrated on full-surround stereoscopic visual design as well as 3D spatial audio to increase immersion in the instrument. Visual calibration has been a key component of this work and we have achieved a seamless view across the multiple projectors lighting the sphere surface. Multi-user interaction using a variety of devices has been another active area of research and is detailed in this document. We believe that all these affordances facilitate immersive, multi-user participation.

The design of the facility is complemented by the development of a computational framework providing an integrated media infrastructure for working with visual, audio, and interactive data. It features a unified programming environment with components for creating interactive, 3D, immersive, multimedia applications that can be scaled from the 3-story instrument to laptops or mobile devices. We found that off-the-shelf VR software and game engines lack the flexibility to represent many forms of complex information (particularly in terms of audio [3]). Media languages such as Max [4] and Processing [5] work well for prototyping, but do not easily scale to large VR simulations. In addition, an in-house, open-source approach was chosen to foster a development community around the facility and to prevent roadblocks in development.

A variety of scientific projects and artistic explorations have driven the design and implementation of the instrument and development framework. We present several of these projects that demonstrate multi-user, multimodal interaction and illustrate our efforts in interactive, immersive data modeling and analysis.

1.1. Related work

The history of unencumbered immersive visualization systems can be traced back to CAVE-like infrastructures designed for immersive VR research [6]. These systems were designed to model virtual reality to real-world problems that allowed a user to move freely in the environment without the need for head-mounted displays and other devices that encumber the user's sense of self [7].

CAVEs had their roots in scientific visualization rather than flight simulation or video games and were closely connected to high performance computing applications [8]. Some of these environments were developed from CAVEs to six-sided cubes as in the StarCAVE [9] and Iowa State's Virtual Reality Application Center. They also developed into multiple-room venues that include immersive theater-like infrastructures, video conferencing rooms, and small immersive working group rooms similar to a small CAVE. Facilities such as these include the Louisiana Immersive Technologies Enterprise (LITE)[2] and Rensselaer Polytechnic's Experimental Media and Performing Arts Center (EMPAC).[3]

As the first VR environments were being designed for a number of varying applications that gravitated toward a single tracked user, smaller more low-cost immersive systems were developed [10–12]. There now exist a plethora of systems from the desktop to plasma screens [13] and large high-resolution displays [14] that allow for immersive visualization in a number of fields. There are also a number of VR laboratories dedicated to specific applications, such as USC's Institute for Creative Technologies, designed for multidisciplinary research focused on exploring and expanding how people engage with computers through virtual characters, video games, simulated scenarios and other forms of human–computer interaction [15] or UC Davis's KeckCAVES (W. M. Keck Center for Active Visualization in the Earth Sciences) [16].

A key difference of the instrument described in this submission to CAVEs and related VR facilities lies in the instrument's ability to provide immersive and interactive surround-view presentations to a group of people[4] who can collaborate with different roles in data navigation and analysis. The screen geometry avoids visual artifacts from sharp discontinuity at corners, enabling seamless immersion even with non-stereoscopic projection, as shown in Fig. 2. Stereo content can be presented to a large set of users who participate in presentations from a bridge through the center of the facility. Users are generally positioned around 5 m distance from the screen, resulting in an audio and stereovision "sweet spot" area that is much larger than in conventional environments.

While we believe that there are many benefits to our instrument design we also acknowledge its limitations. For example, the bridge provides limited room for multiple users to move from one location to another, and so navigation of virtual spaces tends to consist of one user "driving" or "flying" the shared viewpoint with a handheld device, as opposed to, e.g., a (single-user) system based on head-tracking, which could allow navigation in a virtual space via walking, head movements, etc., and would also allow a user to walk all the way around a virtual object to observe it from all sides. Similarly, since every user sees the same left- and right-eye video regardless of location along the bridge, virtual objects closer than the screen appear to track or follow a user as he or she walks along the bridge. This means that correspondence between virtual 3D location (e.g., in an OpenGL scene) and real physical space depends on the viewing position, complicating gestural interaction with virtual objects. Another limitation is that there is almost no ambient light beyond projected content, so cameras used for vision recognition and tracking will be limited to the infrared spectrum. While we do have head tracking capabilities in the instrument, large groups of users are mainly facilitated in non-

[2] http://www.lite3d.com
[3] http://empac.rpi.edu
[4] Groups of up to 30 people can be accommodated. Groups of up to 5 active users are common.